A degree-distance-based connections model with negative and positive externalities

Philipp Moehlmeier, Agnieszka Rusinowska, Emily Tanimura

A degree-distance-based connections model with negative and positive externalities*

Philipp Möhlmeier[1], Agnieszka Rusinowska[2], and Emily Tanimura[3]**

[1] Bielefeld University, BiGSEM, Center for Mathematical Economics, Bielefeld, Germany
philipp.moehlmeier@uni-bielefeld.de

[2] Paris School of Economics - CNRS, Université Paris I Panthéon-Sorbonne, Centre d'Economie de la Sorbonne
agnieszka.rusinowska@univ-paris1.fr

[3] Université Paris I Panthéon-Sorbonne, Centre d'Economie de la Sorbonne
106-112 Bd de l'Hôpital, 75647 Paris, France
Emily.Tanimura@univ-paris1.fr

Abstract. We develop a modification of the connections model by Jackson and Wolinsky (1996) that takes into account negative externalities arising from the connectivity of direct and indirect neighbors, thus combining aspects of the connections model and the co-author model. We consider a general functional form for agents' utility that incorporates both the effects of distance and of neighbors' degree. Consequently, we introduce a framework that can be seen as a degree-distance-based connections model with both negative and positive externalities. Our analysis shows how the introduction of negative externalities modifies certain results about stability and efficiency compared to the original connections model. In particular, we see the emergence of new stable structures, such as a star with links between peripheral nodes. We also identify structures, for example, certain disconnected networks, that are efficient in our model but which could not be efficient in the original connections model. While our results are proved for the general utility function, some of them are illustrated by using a specific functional form of the degree-distance-based utility.

JEL Classification: D85, C70

Keywords: connections model, degree, distance, negative externalities, positive externalities, pairwise stability, efficiency

1 Introduction

The connections model, introduced in the seminal paper of Jackson and Wolinsky (1996) is a setting in which only direct contacts are costly but discounted benefits spill over from indirect neighbors. A natural interpretation is that benefits result from the access to a resource conveyed by the network, such as information or knowledge provided by indirect contacts.

An appealing feature of networks is that they capture the *externalities* that 'occur when the utility of or payoff to an individual is affected by the actions of others, although

* We would like to thank an anonymous referee, the Editor and participants of the conference 'The Role of Externalities in Networks' (Baton Rouge, February 23-24, 2013), in particular, Hans Haller, Ahmed Saber Mahmud and Sudipta Sarangi, and also Herbert Dawid and Tim Hellmann, for helpful comments. Agnieszka Rusinowska and Emily Tanimura acknowledge the support by the National Agency for Research (Agence Nationale de la Recherche), Project DynaMITE (ANR-13-BSH1-0010-01). Philipp Möhlmeier carried out this research within the International Research Training Group Economic Behavior and Interaction Models (EBIM) financed by the German Research Foundation (DFG) under contract GRK 1134/2 and within the Bielefeld Graduate School of Economcis and Management (BiGSEM). Financial support by EBIM, BiGSEM and the Center for Mathematical Economics is gratefully acknowledged.

** Corresponding author.

those actions do not directly involve the individual in question' (Jackson (2008), p. 162). In the connections model, network externalities are positive. An additional link formed by some pair of individuals (weakly) benefits all other agents by providing access to new indirect contacts or by reducing the distance that information has to travel. Such positive aspects of increased connectivity are certainly important. However, in many situations increased connectivity can also have negative side effects. Studying such cases is what motivates the analysis in this paper: we consider a model in which agents benefit from indirect contacts as in the original connections model but in which the connectivity of an agent may also exert a negative externality on his direct and indirect neighbors.

Contexts where this is the case abound. For example, learning of a job opening may be less useful if the information has been communicated to many others. When there is competition for some resource transmitted by the network, the benefits from indirect contacts are reduced when the latter have many connections. In our model, the utility an agent derives from an indirect contact, viewed as the initial sender of an information, is reduced when the latter has a high degree and thus sends the information to many others. However this might not fully account for the negative impact of all other individuals in the communication chain who receive the information. Hence, our model should be viewed only as a simplified or approximate description of the negative effects of connectivity when there is competition for information.

Another negative effect of high connectivity that our model perfectly captures arises because the busyness of agents reduces their availability or productivity. The connections model of Jackson and Wolinsky (1996) could also be interpreted as follows: nodes generate output by themselves but also forward output from others. Now interpret this as a situation in which individuals are involved in projects and generate knowledge by themselves but also receive knowledge from others. Then, a person involved in many projects will have less time to generate output. On the other hand, the more well connected he is, the more knowledge he will receive and forward to his neighborhood. Stated in a provocative way, "well connected people are often great talkers, but networking is time consuming and reduces one's productive time so that the main work is done by others". Nevertheless, the role of such well connected agents is very important: not that they contribute a lot by their own knowledge production, but they provide access to the output of many others.

By integrating the negative effect of the busyness of agents, at first sight, our model looks similar to the well-known co-author model (also introduced in Jackson and Wolinsky (1996)) where the time devoted to a single project decreases with the total number of projects a co-author is involved in. In our version, the nature of the negative externality is similar but information spills over from indirect connections. The co-author model only considers direct collaborations and thus conveys negative effects solely through the busyness of direct co-authors. Hence, our model combines aspects of the connections model and the co-author model. Externalities resulting from additional links can be both positive and negative. New links are useful for reaching indirect partners, but the latter will be more busy, less productive and thus less valuable per se although more efficient as intermediaries. Exploring the tradeoff of these effects and their impact on the stable and efficient network architectures that can arise is the object of our analysis.

A host of papers modify the original connections model in different directions, but most are not directly related to the issues we explore. Generalizing Jackson and Wolinsky (1996), Bloch and Jackson (2007) show that the results of the latter still hold when decay takes a more general functional form. Besides the aforementioned co-author model, the study of negative externalities that is most closely related to ours is the model of Morrill (2011) in which the benefits of a link is a decreasing function of the partner's degree. In fact, our setting generalizes Morrill (2011), where there are no benefits from indirect contacts and externalities are purely negative. Goyal and Joshi (2006) consider a framework where connectivity can generate positive or negative externalities, depending on whether the other agent is a direct, indirect or non-neighbor. They investigate two specific models. The first one captures negative externalities due to overall connectivity and thus addresses situations somewhat different from the ones that motivate our work. In their second model, the marginal benefit of forming a link to some agent is affected both by his and by one's own degree. The authors characterize stable structures, both in cases where the marginal benefits of a link increase with the potential partner's degree and in cases where they decrease as a function of it. Billand et al. (2012) prove the existence of a pairwise stable network in a local spillover game, when the marginal benefit of linking to an agent is decreasing with the degree of the latter and increasing in the own degree. The two aforementioned studies focus on the interplay between an agent's own degree and the degree of his neighbors when spillovers are local, whereas our analysis explores the tradeoff between degree-based negative externalities, reflecting busyness, and the possibility of receiving positive spillovers from distant partners.

We consider a network formation game whose payoffs have a functional form similar to that of the generalized connections model by Bloch and Jackson (2007) – the distance-based utility model, so as to facilitate comparison with the latter, but involve a penalty resulting from the degree of direct and indirect contacts. In this degree-distance-based connections model, we assume a two-variable (instead of one-variable) benefit function which depends on distances to and degrees of (direct and indirect) neighbors.

As in the original connections model, there is multiplicity of pairwise stable structures. We do not give a complete characterization but focus on some cases where outcomes can be compared and contrasted with those of the original connections model. In particular, we analyze stable structures with short diameters. The Jackson and Wolinsky model exhibits two such pairwise stable structures, the star and the complete network with a swift transition from one to the other. In our model, these structures can also be stable, but more interestingly we find new pairwise stable structures with short diameters which could not arise in the original connections model.

The nature of these new structures raises the following question: when direct and indirect contacts are evaluated based on two criteria, their capacity to be intermediaries and to what extent they are available (i.e. not too busy), how should these two roles be distributed? Will we see specialization so that some highly connected agents are valuable only as efficient intermediaries whereas other contacts are counted on for availability, or will we see a more equal distribution of roles where all agents are moderately busy and play a moderate role as intermediaries?

Indeed, we identify two types of architectures both of which ensure short communication paths between all agents but which are organized quite differently. One resembles a star but all peripheral agents also have a "local" neighborhood of direct contacts. In this case, the agents in the network occupy different roles. The center is specialized in the role of intermediary but is too busy to be of much value per se. The agents in the local neighborhood are not useful as intermediaries, but, being less busy, they are valuable in their own right. In the second structure, all agents have similar degrees and contribute equally to providing indirect contacts.

We derive stability conditions for the stable architectures in the original connections model (star/complete/empty) and for the aforementioned new structures. For a given payoff function, these conditions tell us whether one of these structures is stable. However, they provide little intuition for what determines the stability. We pursue our analysis under the additional fairly natural assumption that decay is independent of degree. The concavity of the benefit function with respect to degree is then sufficient to ensure the existence of a range of link costs for which at least one of the new structures with short diameter is stable. More generally, the concavity/convexity of the benefit function with respect to degree is seen to play an important role in determining which structures with short diameters are stable. We also analyze stability and efficiency for extreme levels of decay. For high decay, our stable structures coincide with those in Morrill (2011) which is natural since our model approximates his when decay is large. For small levels of decay, stable structures will be minimally connected with some constraints on degrees.

As shown in Jackson and Wolinsky (1996) strongly efficient networks may not be stable, and conversely networks need not be efficient even when they are uniquely stable. Buechel and Hellmann (2012) show that inefficient outcomes can be related to the nature of the externalities. They introduce the notion of over-connected (under-connected) networks – those which can be socially improved by the deletion (addition) of links. The authors prove that for positive externalities no stable network can be over-connected. Negative externalities tend to induce over-connected networks, and under some additional conditions, no stable network can be under-connected. In our model, the new structures with short diameters, while stable, would typically not be efficient when the network is large. The same is true for the complete graph. We show this without actually characterizing *the* efficient network. Finally, we show that under certain conditions the star will be uniquely efficient. The conditions required in our proof are quite restrictive but compatible with the stability of the star structures with peripheral links. Thus, our model can indeed generate over-connectedness as defined by Buechel and Hellmann (2012), which could never occur in the original connections model.

The rest of the paper is structured as follows. In Section 2 we recapitulate some preliminaries on networks, the Jackson and Wolinsky connections model and existing extensions. In Section 3 we present our model. Pairwise stability and efficiency are studied in Sections 4 and 5, respectively. We begin by providing some illustrating examples in networks of small size and then turn to the general analysis of stability and efficiency. In Section 6 we mention some possible extensions. Some proofs are presented in the Appendix.

2 The connections model and its modifications

In this section we first present the preliminaries on networks (see, e.g., Jackson and Wolinsky (1996); Jackson (2008)) and then briefly recapitulate some models related to our work: the connections model and the co-author model of Jackson and Wolinsky (1996), the distance-based model by Bloch and Jackson (2007), and the model with degree-based utility functions by Morrill (2011).

Let $N = \{1, 2, \ldots, n\}$ denote the set of players (agents). A *network* g is a set of pairs $\{i, j\}$ denoted for convenience by ij, with $i, j \in N$, $i \neq j$,[1] where ij indicates the presence of a pairwise relationship and is referred to as a *link* between players i and j. Nodes i and j are directly connected if and only if $ij \in g$.

A *degree* $\eta_i(g)$ of agent i counts the number of links i has in g, i.e.,

$$\eta_i(g) = |\{j \in N \mid ij \in g\}|$$

We can identify two particular network relationships among players in N: the *empty network* $g^{\emptyset}$ without any link between players, and the *complete network* g^N which is the set of all subsets of N of size 2. The set of all possible networks g on N is $G := \{g | g \subseteq g^N\}$.

By $g + ij$ ($g - ij$, respectively) we denote the network obtained by adding link ij to g (deleting link ij from g, respectively). Furthermore, by g_{-i} we denote the network obtained by deleting player i and all his links from the network g.

Let $N(g)$ ($n(g)$, respectively) denote the set (the number, respectively) of players in N with at least one link, i.e., $N(g) = \{i \mid \exists j \text{ s.t. } ij \in g\}$ and $n(g) = |N(g)|$.

A *path* in g connecting i_1 and i_K is a set of distinct nodes $\{i_1, i_2 \ldots, i_K\} \subseteq N(g)$ such that $\{i_1 i_2, i_2 i_3, \ldots, i_{K-1} i_K\} \subseteq g$.

A network g is *connected* if there is a path between any two nodes in g.

The network $g' \subseteq g$ is a *component* of g if for all $i \in N(g')$ and $j \in N(g')$, $i \neq j$, there exists a path in g' connecting i and j, and for any $i \in N(g')$ and $j \in N(g)$, $ij \in g$ implies that $ij \in g'$.

A *star* g^* is a connected network in which there exists some node i (referred to as the *center* of the star) such that every link in the network involves node i.

The *value* of a graph is represented by $v : G \to \mathbb{R}$. By V we denote the set of all such functions. In what follows we will assume that the value of a graph is an aggregate of individual utilities, i.e., $v(g) = \sum_{i \in N} u_i(g)$, where $u_i : G \to \mathbb{R}$.

A network $g \subseteq g^N$ is *strongly efficient (SE)* if $v(g) \geq v(g')$ for all $g' \subseteq g^N$.

A network $g \in G$ is *pairwise stable (PS)* if:

(i) $\forall\ ij \in g$, $u_i(g) \geq u_i(g - ij)$ and $u_j(g) \geq u_j(g - ij)$ and
(ii) $\forall\ ij \notin g$, if $u_i(g) < u_i(g + ij)$ then $u_j(g) > u_j(g + ij)$.

In the symmetric *connections model* by Jackson and Wolinsky (1996), the utility of each player i from network g is defined as

$$u_i^{JW}(g) = \sum_{j \neq i} \delta^{l_{ij}(g)} - c\eta_i(g) \tag{1}$$

where $0 < \delta < 1$ denotes the undiscounted valuation of a connection, $l_{ij}(g)$ denotes the distance between i and j in terms of the number of links in the shortest path between them in g (with $l_{ij}(g) = \infty$, if there is no path connecting i and j in g) and $c > 0$ determines the costs for a direct connection.

Jackson and Wolinsky (1996) (Proposition 1) show that the complete, the empty or the star graph can be uniquely strongly efficient (depending on c and δ). More precisely, they prove that the unique SE network in the symmetric connections model is:

[1] Loop ii is not a possibility in this setting.

(i) the complete network g^N if $c < \delta - \delta^2$

(ii) a star g^* if $\delta - \delta^2 < c < \delta + \frac{(n-2)\delta^2}{2}$

(iii) no links if $\delta + \frac{(n-2)\delta^2}{2} < c$.

They also examine pairwise stability in the symmetric connection model. By virtue of Jackson and Wolinsky (1996) (Proposition 2), in the symmetric connections model:

(i) A pairwise stable graph has at most one (non-empty) component.

(ii) For $c < \delta - \delta^2$, the unique PS network is the complete graph g^N.

(iii) For $\delta - \delta^2 < c < \delta$, a star g^* encompassing all players is PS, but not necessarily the unique PS graph.

(iv) For $\delta < c$, any PS network which is non-empty is such that every player has at least two links (and thus is inefficient).

Jackson and Wolinsky (1996) also present the *co-author model*, in which nodes are interpreted as researchers and a link represents a collaboration between two researchers. The utility function of each player i in network g is given by

$$u_i^{co}(g) = \sum_{j:ij \in g} w_i(n_i, j, n_j) - c(n_i) \tag{2}$$

where $w_i(n_i, j, n_j)$ is the utility of i derived from a link with j when i and j are involved in n_i and n_j projects, respectively, and $c(n_i)$ is the cost to i of maintaining n_i links.

Bloch and Jackson (2007) introduce an extension of the original connections model – the distance-based model, where the utility of agent i is given by

$$u_i^{dist}(g) = \sum_{j \neq i} f(l_{ij}(g)) - c\eta_i(g) \tag{3}$$

with f nonincreasing in $l_{ij}(g)$; see also Jackson (2008) for the presentation of this distance-based model.

Morrill (2011) models situations in which any new relationship causes negative externalities. The payoff of each player from a link is a decreasing function of the number of links maintained by his partner. A utility function is *degree-based* if there exists a decreasing function ϕ such that

$$u_i^{deg}(g) = \sum_{j:ij \in g} \phi(\eta_j(g)) - c\eta_i(g) \tag{4}$$

3 A degree-distance-based connections model

In Jackson and Wolinsky (1996), an additional link induces only positive externalities. We suggest a modification that also generates negative externalities due to increasing connectivity. Every agent benefits from his direct and indirect connections, but it is additionally assumed that the higher the degree of a direct or indirect partner, the less valuable is this connection. In order to remain close to the connections model we follow existing generalizations of Jackson and Wolinsky (1996), by considering the utility of agent i given by

$$\widetilde{u}_i(g) = \sum_{j \neq i} b(l_{ij}(g), \eta_j(g)) - c\eta_i(g) \tag{5}$$

where $b : \{1, ..., n-1\}^2 \to \mathbb{R}^+$ is the net benefit that an agent receives from the direct and indirect connections, and $c > 0$ is the cost for a direct connection. It is assumed that for all $l_{ij}(g)$, $b(l_{ij}(g), k)$ is nonincreasing in degree k, and for all $\eta_j(g)$, $b(l, \eta_j(g))$ is nonincreasing in distance l. Moreover, if there is no path connecting i and j in g, i.e., if $l_{ij}(g) = \infty$, then we set $b(\infty, \eta_j) = 0$ for every $\eta_j \in \{0, 1, ..., n-1\}$. In particular, $\widetilde{u}_i(g^\emptyset) = 0$ for every $i \in N$.

In the original connections model the benefit term is expressed using a single parameter δ. If we expressed the benefit function in our model using parameters that regulate decay with distance and utility loss due to degree, we could write

$$b(l+1, \eta_j(g)) = \delta_{l,\eta_j(g)} b(l, \eta_j(g)), \quad b(l, \eta_j(g)+1) = c_{l,\eta_j(g)} b(l, \eta_j(g))$$

where $\delta_{l,\eta_j(g)} \in (0,1)$ expresses the decay between distance l and $(l+1)$ for a fixed degree $\eta_j(g)$, and $c_{l,\eta_j(g)} \in (0,1)$ expresses the utility loss due to an additional link increasing the degree from $\eta_j(g)$ to $(\eta_j(g)+1)$ for a fixed distance l. This gives much versatility, in particular, decay does not need to be constant with distance.

Since we aim to analyze negative externalities resulting from the connectivity of direct and indirect neighbors, we will assume that the benefit function b is decreasing in degree (and in distance), except when mentioning explicitly the original connections model as a particular case of the degree-distance-based model.

Our framework also generalizes the degree-based model by Morrill (2011). We have

$$\phi(\eta_j(g)) = b(1, \eta_j(g)), \text{ for all } \eta_j(g) \in \{1, ..., n-1\} \tag{6}$$

The generalized model defined by (5) also extends the distance-based model considered in Bloch and Jackson (2007) and recapitulated in (3). In other words, we consider a two-variable (instead of one-variable) benefit function.

In some examples, we will use a specific form of the degree-distance-based utility, which is very close to the original connections model, except that it incorporates an additional information about the degree of direct and indirect neighbors. More precisely, to illustrate some of our results, the following utility of agent i will be used:

$$u_i(g) = \sum_{j \neq i} \frac{1}{1 + \eta_j(g)} \delta^{l_{ij}(g)} - c\eta_i(g) \tag{7}$$

that is, we will set $b(l_{ij}(g), \eta_j(g)) = \frac{1}{1+\eta_j(g)} \delta^{l_{ij}(g)}$.

An idea somewhat similar to the one expressed by our model is presented in Haller (2012) who studies a non-cooperative model of network formation. He considers two examples with negative network externalities in which the values of information are endogenously determined and depend on the network. This is in line with the idea that it is harder to access the information from an agent who has more direct neighbors.

4 Pairwise stability in the model

4.1 Stability of the star, the complete graph and the empty graph

Next, we examine pairwise stability (PS) in the model. In order to compare results in our model with those of Jackson and Wolinsky (1996), we start by analyzing the stability of the architectures which were prominent there: the empty network, the star and the complete graph. Furthermore, we check if a non-empty PS network must be connected. In the connections model of Jackson and Wolinsky (1996), any pairwise stable graph has at most one (non-empty) component. We will show that it is not necessarily the case in our model. After establishing the stability conditions for $g^{\emptyset}$, g^* and g^N, we will look for other PS structures. We will begin by considering networks in which all agents are at distance at most 2 from each other. In such a network, the benefit of adding a link to an agent of degree $k-1$ is

$$\widetilde{f}(k) := b(1,k) - b(2,k-1) \text{ for } k \in \{2,...,n-1\}, \text{ and } \widetilde{f}(1) := b(1,1) \tag{8}$$

Note that in the original connections model $\widetilde{f}(k) = \delta - \delta^2$ for each k. We have the following results on pairwise stability of the three prominent architectures.

Proposition 1 *In the degree-distance-based connections model defined by (5):*

(i) The empty network $g^{\emptyset}$ is PS if and only if $\widetilde{f}(1) \leq c$.
(ii) The star g^* with $n \geq 3$ encompassing all players is PS if and only if

$$\widetilde{f}(2) \leq c \leq \min\Big(\widetilde{f}(1), b(1,n-1) + (n-2)b(2,1)\Big) \tag{9}$$

This cost range is non empty whenever $\widetilde{f}(2) \leq b(1,n-1) + (n-2)b(2,1)$.
(iii) The complete network g^N with $n \geq 3$ is PS if and only if

$$c \leq \widetilde{f}(n-1) \tag{10}$$

(iv) The unique PS network is the complete network g^N if

$$c < \min_{1 \leq \eta_k \leq n-2} \widetilde{f}(\eta_k + 1) \tag{11}$$

(v) g^* and g^N are simultaneously PS if and only if $\widetilde{f}(2) \leq c \leq \widetilde{f}(n-1)$. This cost range is non empty whenever $\widetilde{f}(2) \leq \widetilde{f}(n-1)$. In particular, if $\widetilde{f}(n-1) < c < \widetilde{f}(2)$, then neither the complete graph nor the star is PS.
(vi) A PS network may have more than one (non-empty) component.

Proof: (i) Consider any two agents $i, j \in g^{\emptyset}$. $\widetilde{u}_i(g^{\emptyset}+ij) - \widetilde{u}_i(g^{\emptyset}) = \widetilde{u}_j(g^{\emptyset}+ij) - \widetilde{u}_j(g^{\emptyset}) = b(1,1) - c = \widetilde{f}(1) - c$. Hence, if $\widetilde{f}(1) > c$, then both players profit from establishing the link, and therefore $g^{\emptyset}$ is not PS. If $\widetilde{f}(1) \leq c$, then $\widetilde{u}_i(g^{\emptyset}+ij) - \widetilde{u}_i(g^{\emptyset}) \leq 0$ and $\widetilde{u}_j(g^{\emptyset}+ij) - \widetilde{u}_j(g^{\emptyset}) \leq 0$ which means that $g^{\emptyset}$ is PS.
(ii) Consider the star g^* with $n \geq 3$ agents. Take the center of the star i and two arbitrary agents j, k, where $j \neq i$, $k \neq i$, and $j \neq k$. This means that $ij \in g^*$ but $jk \notin g^*$. For stability the following conditions must hold:
(A) $\widetilde{u}_i(g^*) - \widetilde{u}_i(g^* \setminus ij) \geq 0$ and (B) $\widetilde{u}_j(g^*) - \widetilde{u}_j(g^* \setminus ij) \geq 0$ and (C) $\widetilde{u}_j(g^*+jk) - \widetilde{u}_j(g^*) \leq 0$.
(A): $\widetilde{u}_i(g^*) - \widetilde{u}_i(g^* \setminus ij) = b(1,1) - c = \widetilde{f}(1) - c$. Hence, (A) holds iff $\widetilde{f}(1) \geq c$.
(B): $\widetilde{u}_j(g^*) - \widetilde{u}_j(g^* \setminus ij) = b(1,n-1) + (n-2)b(2,1) - c$. Hence, (B) holds iff $b(1,n-1) + (n-2)b(2,1) \geq c$.
(C): $\widetilde{u}_j(g^*+jk) - \widetilde{u}_j(g^*) = b(1,2) - b(2,1) - c = \widetilde{f}(2) - c$. Hence, (C) holds iff $\widetilde{f}(2) \leq c$.

Hence, we get $\widetilde{f}(2) \leq c \leq \min\left(\widetilde{f}(1), b(1, n-1) + (n-2)b(2,1)\right)$.

(iii) Let $n \geq 3$. Consider any two agents $i, j \in g^N$. We have $\widetilde{u}_i(g^N) - \widetilde{u}_i(g^N - ij) = \widetilde{u}_j(g^N) - \widetilde{u}_j(g^N - ij) = b(1, n-1) - b(2, n-2) - c = \widetilde{f}(n-1) - c$.

(iv) Consider two arbitrary agents i and j, $j \neq i$ such that $ij \notin g$, $\eta_i > 0$ and $\eta_j > 0$. Then we have $\widetilde{u}_i(g + ij) - \widetilde{u}_i(g) \geq b(1, \eta_j + 1) - b(2, \eta_j) - c = \widetilde{f}(\eta_j + 1) - c$ and $\widetilde{u}_j(g + ij) - \widetilde{u}_j(g) \geq \widetilde{f}(\eta_i + 1) - c$. If $\eta_i \eta_j = 0$, then $\widetilde{u}_i(g + ij) - \widetilde{u}_i(g) \geq b(1, \eta_j + 1) - c$ and $\widetilde{u}_j(g + ij) - \widetilde{u}_j(g) \geq b(1, \eta_i + 1) - c$. Hence, if $c < \min_{1 \leq \eta_k \leq n-2} \widetilde{f}(\eta_k + 1)$, then any two agents who are not directly connected benefit from forming a link.

(v) We have $\widetilde{f}(n-1) = b(1, n-1) - b(2, n-2) < b(1, n-1) + (n-2)b(2,1)$. Moreover, from the nonincreasingness of b in degree, $\widetilde{f}(n-1) < b(1,1) = \widetilde{f}(1)$. Hence, from (9) and (10), g^* and g^N are simultaneously PS if and only if $\widetilde{f}(2) \leq c \leq \widetilde{f}(n-1)$.

(vi) The general existence of pairwise stable disconnected structures is given in Proposition 8. In small networks we can also find other types of architectures. Consider for example g given in Figure 1 and the utility function given by (7).

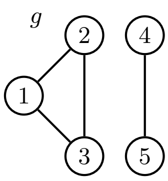

Fig. 1. A PS network with two components in the degree-distance-based connections model

Network g is PS if $\frac{1}{4}\delta + \frac{2}{3}\delta^2 < c \leq \frac{1}{3}\delta - \frac{1}{2}\delta^2$, e.g., for $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$. Since we have two groups of symmetric agents $(1, 2, 3$ and $4, 5)$, we only need to calculate the following:
$u_2(g) - u_2(g \setminus 23) = \frac{2}{3}\delta - 2c - \left(\frac{1}{3}\delta + \frac{1}{2}\delta^2 - c\right) = \frac{1}{3}\delta - \frac{1}{2}\delta^2 - c \geq 0$
$u_4(g) - u_4(g \setminus 45) = \frac{1}{2}\delta - c \geq 0$
$u_2(g + 24) - u_2(g) = \frac{1}{3}\delta + \frac{1}{2}\delta^2 - c$ and $u_4(g + 24) - u_4(g) = \frac{1}{4}\delta + \frac{2}{3}\delta^2 - c$
Note that if $\frac{1}{3}\delta - \frac{1}{2}\delta^2 - c \geq 0$, then $\frac{1}{2}\delta - c > 0$ and $\frac{1}{3}\delta + \frac{1}{2}\delta^2 - c > 0$. Hence, g will be PS if $\frac{1}{3}\delta - \frac{1}{2}\delta^2 \geq c$ and $\frac{1}{4}\delta + \frac{2}{3}\delta^2 < c$. ■

Note that Proposition 1 confirms, in particular, the results on pairwise stability of g^N, g^* and $g^\emptyset$ in the Jackson and Wolinsky model. Assume now that the benefit function b is strictly decreasing in degree – for simplicity, take the degree-distance-based model given by (7). Naturally, since $\delta - \delta^2 > \frac{\delta}{n} - \frac{\delta^2}{n-1}$ for any $n \geq 2$, if g^N is PS in our model, then it is also PS in the Jackson and Wolinsky framework under the same parameters. Moreover, if $g^\emptyset$ is PS in the original connections model, then it is also PS in our framework. Roughly speaking, the costs under which the star g^* is PS in our model are rather lower than the costs under which g^* is PS in the Jackson and Wolinsky model for the same δ and n. For $\delta < \frac{1}{2}$, g^* cannot be PS in both frameworks at the same time, but for $\delta \geq \frac{1}{2}$ such an overlap of costs under which the star is PS is non-empty[2].

[2] In the Jackson and Wolinsky model, g^* is PS if $\delta - \delta^2 < c < \delta$. In our model (7), g^* is PS if $\frac{\delta}{3} - \frac{\delta^2}{2} \leq c \leq \min\left(\frac{\delta}{2}, \frac{\delta}{n} + \frac{(n-2)\delta^2}{2}\right)$. As $\frac{\delta}{3} - \frac{\delta^2}{2} < \delta - \delta^2$, the costs range under which g^* is PS in both frameworks is

Proposition 1(v) shows the existence of a region where the star and the complete graph are simultaneously PS. This could never occur in the original connections model where the regions of stability for these two structures were disjoint. However, we should note that, for instance, for the degree-distance-based utility given by (7), $\lim_{n\to\infty}\widetilde{f}(n-1)=0$ so that the possible cost range for which the complete network and the star are simultaneously PS is very small in large networks.

Figure 2 (left) illustrates the pairwise stability regions for the three simple architectures for the model given by (7) with $n=9$, $\delta\in(0,1)$ and $c\in(0,0.5]$[3]. In this figure, the green area indicates the stability region of $g^{\emptyset}$, the red area the one for g^N and the yellow area the one for g^*. The overlapping (quite small) orange area indicates the parameter region in which g^N and g^* are simultaneously PS, and the white area in which none of the three simple structures are PS.

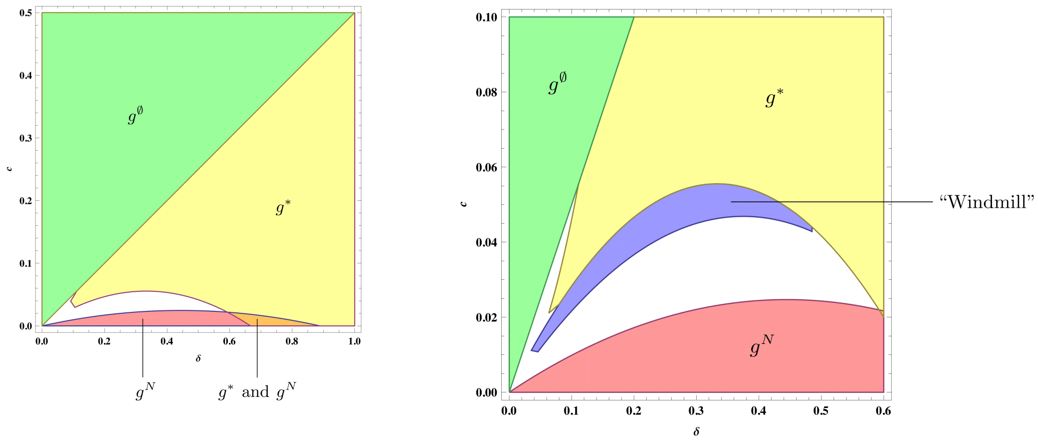

Fig. 2. PS regions in the degree-distance-based connections model given by (7) ($n=9$): only $g^{\emptyset}$ (green area), only g^* (yellow area), only g^N (red area), none of these three (white area); *Left* – g^* and g^N simultaneously (orange area); *Right* – "windmill" (blue region)

4.2 Other pairwise stable structures in the degree-distance-based model: examples and illustration

Next, we will be interested in PS architectures of the degree-distance-based connections model, other than those analyzed in the previous section. An example of a PS structure that can occur in the white area in Figure 2 (left) is the "windmill" structure shown in Figure 3. This is a specific example of what we will call a *core-periphery structure.*

Definition 1 *In a core periphery structure with periphery degree η_m, one node, the center, is linked to all other nodes, and every node other than the center has the same degree η_m.*

$\delta-\delta^2\le c\le\min\left(\frac{\delta}{2},\frac{\delta}{n}+\frac{(n-2)\delta^2}{2}\right)$. Note that $\delta-\delta^2\le\frac{\delta}{2}$ if and only if $\delta\ge\frac{1}{2}$, and therefore g^* cannot be PS if $\delta<\frac{1}{2}$. For $\delta\ge\frac{1}{2}$, the overlap of costs is non-empty, as $\delta-\delta^2\le\min\left(\frac{\delta}{2},\frac{\delta}{n}+\frac{(n-2)\delta^2}{2}\right)$.

[3] The calculations have been done in *Mathematica*. The details can be provided upon request.

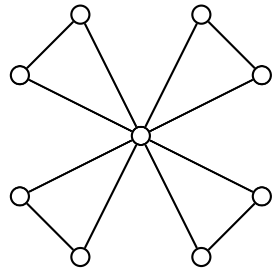

Fig. 3. "Windmill" as an example of a PS network in the degree-distance-based connections model

To get a more precise feeling for the range of parameters in which such a "windmill" structure is stable, consider Figure 2 (right). Compared to Figure 2 (left), Figure 2 (right) is zoomed in and shows an additional blue area in which the "windmill " structure is PS. The green, yellow, red and white areas have the same meaning as before.

Figure 4 presents all PS networks for $n = 3$ with the corresponding parameters for the model (7). Note that a network with one link can be PS in our framework (with b being strictly decreasing in degree), contrary to the Jackson and Wolinsky model. Figure 4 confirms nicely Proposition 1(v). In particular, g^* and g^N are simultaneously PS only for $c = \frac{\delta}{3} - \frac{\delta^2}{2}$.

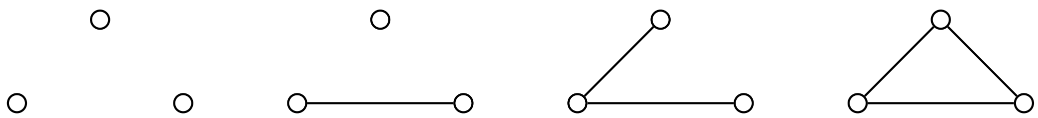

Fig. 4. PS networks in the degree-distance-based connections model given by (7) (from left to right): (i) $\delta \leq 2c$, (ii) $0 < \delta < \frac{1}{3}$ and $\frac{\delta}{3} + \frac{\delta^2}{2} < c \leq \frac{\delta}{2}$, (iii) ($0 < \delta < \frac{1}{3}$, $\frac{\delta}{3} - \frac{\delta^2}{2} \leq c \leq \frac{\delta}{3} + \frac{\delta^2}{2}$) or ($\frac{1}{3} \leq \delta < 1$, $\frac{\delta}{3} - \frac{\delta^2}{2} \leq c \leq \frac{\delta}{2}$), (iv) $c \leq \frac{\delta}{3} - \frac{\delta^2}{2}$

Let us now illustrate, for some small network sizes, examples of other PS structures that can appear. In some cases, these architectures are not stable in the original connections model which make them interesting per se. Many of the architectures that we see in these examples can also be shown to exist in a network of arbitrary size n in some parameter range, as will be shown in a later section. Figures 5 and 6 show some examples of different PS structures for $n = 5$ and $n = 6$, respectively, for the model given by (7) with $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$. From among these examples, only the two regular networks (with the degree $\eta_i = 2$ for every $i \in N$) can be PS under some parameters in the original connections model. The remaining networks which are PS in our framework could never be PS in the original connections model. Note that four of these networks contain two components. In Figure 5 these are the first network (on the left) that has been used in the proof of Proposition 1(vi), and the second network with one isolated player and four players, each having the degree equal to 2. In Figure 6 these are the first and the second network, both having two non-empty components.

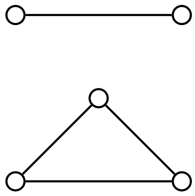

Fig. 5. Some PS networks in the degree-distance-based connections model given by (7) for $n = 5$, $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$

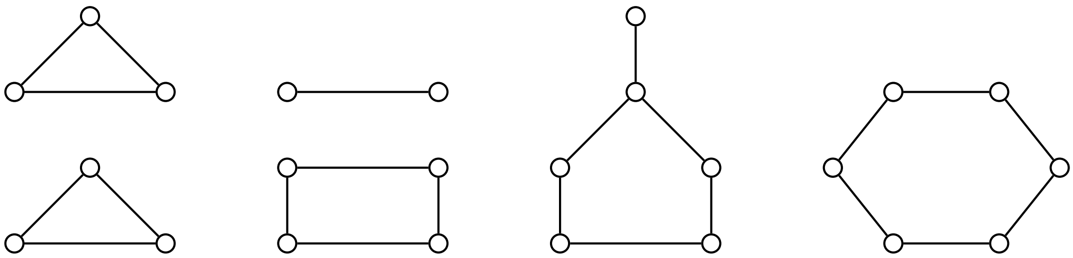

Fig. 6. Some PS networks in the degree-distance-based connections model given by (7) for $n = 6$, $\delta = \frac{1}{15}$ and $c = \frac{107}{5400}$

4.3 Stability analysis for new architectures with short diameters

In the previous sections, we provided examples of networks that are PS in our framework for small values of n. These illustrate some of the general results about architectures that can be PS for an arbitrary size n for some values of the decay and cost parameters. In contrast, most of them can never be PS in the original connections model. We have already seen that there exist PS structures other than the star and the complete graph, in particular, ones that are intermediary between these two in the sense of link inclusion, that is g is PS and $g^* \subset g \subset g^N$. This is never possible in the original Jackson and Wolinsky model. Figure 3 depicted an example of such a structure, the "Windmill" which is an example of a PS structure that contains a star but which also comprises links between the nodes in the periphery. We prove the following result about the stability of stars with peripheral links.

Proposition 2 *Assume that the benefit b and the cost c are such that there exists an $\eta_m \in N$ such that $\widetilde{f}(\eta_m + 1) < c \leq \widetilde{f}(\eta_m)$. Assume that moreover, n is such that $\widetilde{f}(n - 1) + (n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2))\,(b(2, \eta_m) - b(3, \eta_m)) > c$. Then any core-periphery structure with periphery degree η_m is PS.*

Proof: Let us show that under the conditions stated the described structure is pairwise stable. Every peripheral node has degree η_m. Breaking the link to another peripheral node does not modify the benefits from nodes that can be reached at distance 2. Only direct benefits are lost. None of those peripheral nodes wants to break a local link with another peripheral agent of degree η_m because $\widetilde{u}_j(g) - \widetilde{u}_j(g - lj) = b(1, \eta_m) - b(2, \eta_m - 1) - c = \widetilde{f}(\eta_m) - c \geq 0$ by assumption. No agent wants to add a link to a peripheral agent whose degree in g is η_m because doing this would decrease the utility, as $\widetilde{u}_j(g + jm) - \widetilde{u}_j(g) = b(1, \eta_m + 1) - b(2, \eta_m) - c = \widetilde{f}(\eta_m + 1) - c < 0$ by assumption.

Every peripheral node has an incentive to maintain a link to the center for the following reason: Without a link to the center, an agent can reach at most $(\eta_m - 1)$ neighbors. Each neighbor has $(\eta_m - 2)$ links to nodes different from the center and the respective agent. Hence, without a link to the center, an agent can reach at most $(\eta_m - 1)(\eta_m - 2) + 1$ nodes at distance 2 (including one indirect link to the center). To identify the nodes who are at distance 3 in absence of a link to the center, one has to subtract from the overall number of nodes n the agent himself, his direct neighbors and the neighbors of degree 2. Consequently, by breaking a link to the center, at least $n-1-(\eta_m-1)-((\eta_m - 1)(\eta_m - 2) + 1) = n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2)$ nodes move to distance 3. Thus $\widetilde{u}_j(g) - \widetilde{u}_j(g - ij) \geq b(1, n - 1) - b(2, n - 2) + (n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2)) (b(2, \eta_m) - b(3, \eta_m)) - c = \widetilde{f}(n - 1) + (n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2)) (b(2, \eta_m) - b(3, \eta_m)) - c$.
We show that for given b and c, we can find n such that

$$\widetilde{f}(n - 1) + (n - 1 - \eta_m - (\eta_m - 1)(\eta_m - 2)) (b(2, \eta_m) - b(3, \eta_m)) - c > 0 \tag{12}$$

Note that $b(2, \eta_m) > b(3, \eta_m)$. Moreover, $b(1, k) \geq 0$ for all k, and therefore from (8) we have $\widetilde{f}(n-1) \geq -b(2, n-2) \geq -b(2, 1)$. Hence, $c - \widetilde{f}(n-1) \leq c + b(2, 1)$ and then every n such that $n > \frac{c+b(2,1)}{b(2,\eta_m)-b(3,\eta_m)} + 1 + \eta_m + (\eta_m - 1)(\eta_m - 2)$ satisfies (12). ■

Note that in the original connections model, where $\widetilde{f}$ is constant, the first assumption of Proposition 2 which requires a decreasing $\widetilde{f}$ is never satisfied.

The stable structure with inhomogeneous degrees given in Proposition 2 illustrates the fact that in the degree-distance based model, a node can be attractive either because it is not too "busy" with other neighbors or because it is highly connected and provides indirect benefits. We also have examples of stable structures where the degree distribution is essentially homogeneous.

Proposition 3 *Let g be a network such that $l_{ij}(g) \leq 2$ for all i, j, and there are at least $(n - 1)$ nodes with identical degree k and at most one node j such that $\eta_j < k$. Then g is PS in the cost range $\widetilde{f}(k + 1) < c < \widetilde{f}(k)$ and $\widetilde{f}(\eta_j) > c$. In particular, for the utility function given by (7), this cost range is non empty if and only if $k > \frac{\delta+\sqrt{\delta}}{1-\delta}$ and $\widetilde{f}(\eta_j) > c$.*

Proof: No agent wants to add a link to an agent with degree k since $\widetilde{f}(k + 1) < c$. Moreover, no agent wants to delete a link to an agent with degree k since the loss of utility is at least $\widetilde{f}(k)$ which exceeds the saving of c. Note that there may be a single agent j such that $\eta_j < k$. This agent does not want to form a link to any other agent since all other agents have degree k. Moreover, since $l_{ij} \leq 2$, the only path that is shortened is the one to the player one links to, and this path is reduced by one link. No agent wants to drop a link to agent j since $\widetilde{f}(\eta_j) > c$. For the utility function given by (7), there exists a cost c such that $\widetilde{f}(k+1) < c < \widetilde{f}(k)$ if k belongs to the interval where $\widetilde{f}(k)$ is decreasing, i.e., on $[\frac{\delta+\sqrt{\delta}}{1-\delta}, n - 1]$. ■

Structures satisfying the properties stated in Proposition 3 clearly exist. We can give some (non exhaustive) examples.

Proposition 4 *Pairwise stable networks with equal degree: For n, M and l such that $M = \sqrt{n} \in N$, and l is a divisor of M, there exists a pairwise stable network g such that $l_{ij}(g) \leq 2$ for all i, j and all nodes have a degree equal to $k = (M-1) + (M-1)l$. It consists of M completely connected components or islands. Each node is linked to exactly l nodes on all other islands.*

Proof: Divide $n = M^2$ into M disjoint sets of size M, which we index by $m = 1, ..., M$. Divide each set of M agents into disjoint sets of size l $(S_i^m)_{i=1}^{i=M/l}$. Link agents from distinct islands if they have the same i. This ensures that each agent has exactly $M-1+(M-1)l$ neighbors. Moreover, $l_{ij}(g) \leq 2$ for all i, j. ■

We note that since $k = (M-1) + (M-1)l \geq 2(\sqrt{n} - 1)$, the structures defined above only exist for fairly large degrees. We also note that whenever $l > 1$, the resulting homogeneous network is such that the removal of a single link does not change the diameter.

Architectures with small diameters: further interpretation and analysis of the stability conditions We established conditions for the stability of the main PS structures considered in Jackson and Wolinsky (1996) as well as other structures with small diameters: windmill structures and equal degree island models. However, these conditions did not provide much intuition for what really lies behind the stability of the various structures.

The conditions did show that the stability of some of the structures was only compatible with certain behaviors of the function $\widetilde{f}$. For example, Propositions 2, 3 require $\widetilde{f}$ to be decreasing. In this section we show that, under an additional assumption, namely that decay is independent of degree, the behavior of the function $\widetilde{f}$ (increasing or decreasing) can be related to the convexity/concavity of the benefit with respect to degree. Conditions related to concavity/convexity are easy to interpret and the results in this section should make it apparent that the conditions ensuring the stability of the new structures with short diameters analyzed in the previous section are not difficult to satisfy.

Throughout this section, we make the assumption that decay of the benefit is independent of the degree of the neighbor in the sense that $\frac{b(l+1,k)}{b(l,k)} := \delta_l$ for every k, that is, the decay of the benefit may vary with the distance l but not with the degree k.

The following proposition links the behavior of the function $\widetilde{f}$ to the convexity/concavity of the benefit with respect to degree.

Proposition 5 *We have the following:*

- *$\widetilde{f}(k)$ is decreasing whenever $b(l, k)$ is concave in the degree k.*
- *If $b(l, k)$ is convex in the degree k, then there exists a level of decay $0 < \delta_m < 1$ such that $\widetilde{f}(k)$ is decreasing whenever $\delta_1 \leq \delta_m$, and there exists a $0 < \delta_M < 1$ such that $\widetilde{f}(k)$ is increasing whenever $\delta_1 \geq \delta_M$.*

Proof: Consider $\widetilde{f}(k) - \widetilde{f}(k+1) = b(1,k) - b(2,k-1) - (b(1,k+1) - b(2,k)) = b(1,k) - b(1,k+1) - \delta_1 (b(1,k-1) - b(1,k))$. This quantity is positive if and only if $\frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)} \geq \delta_1$. If b is concave in degree, then $\frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)} \geq 1 \geq \delta_1$, so that $\widetilde{f}(k)$ will always be decreasing. Now, if b is convex, then $0 < \frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)} < 1$ for all $k \in 1, ..., n-1$. Define $\delta_M =: \max_k \frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)}$ and $\delta_m =: \min_k \frac{b(1,k)-b(1,k+1)}{b(1,k-1)-b(1,k)}$, then $\widetilde{f}(k)$ is decreasing for $\delta_1 \leq \delta_m$ and increasing when $\delta_1 \geq \delta_M$. ■

Consequently, the conditions in Propositions 2 and 3 are compatible with a benefit function that is concave in degree and with one that is convex in degree only if decay is large. Other results, such as the existence of simultaneous stability of the star and the complete network (see Proposition 1) cannot occur for a b concave in degree, but are compatible with a b convex in degree, for which the function $\widetilde{f}$ exhibits a wider range of behaviors.

The behavior of the function $\widetilde{f}$ is crucial for selecting the stable structures with diameter 2. We will now return to the stability conditions of the structures considered in the previous section. When $\widetilde{f}$ is increasing, we have a precise characterization of the stable networks with diameter 2 which have the property that the removal of a single link does not increase the diameter.

Proposition 6 *If $\widetilde{f}$ is increasing, then:*

- *there is a cost range for which the star is stable*
- *there is a cost range for which the complete graph is stable*
- *there is a cost range for which the star and the complete graph are stable*
- *a stable network g such that $l_{ij}(g-kl) \leq 2$ for all i,j and all kl must be a complete network.*

Proof: The function $\widetilde{f}$ is increasing on $[2, n-1]$. However, $\widetilde{f}(1) = b(1,1) > \widetilde{f}(2), ..., \widetilde{f}(1) > \widetilde{f}(n-1)$. The complete network is stable if $c < \widetilde{f}(n-1)$. The star is stable if $\widetilde{f}(2) < c < \widetilde{f}(1)$ and if $\widetilde{f}(n-1) + (n-2)b(2,1) > c$. The condition $\widetilde{f}(2) < c < \widetilde{f}(n-1) < \widetilde{f}(1)$ is sufficient for simultaneous stability of the star and the complete network.

For the last property, consider a stable network g such that $l_{ij}(g-kl) \leq 2$ for all i,j and all kl. The existence of a node of degree 1 contradicts the property $l_{ij}(g-kl) \leq 2$ for all i,j and all kl, because the removal of such a node disconnects the network. First note that if there are two nodes in g whose degrees are different from $n-1$, these nodes must be directly linked. Indeed, suppose that their degrees are η and η'. Neighbors of these agents do not lose any indirect benefits by breaking with them. Hence, stability requires that $\widetilde{f}(\eta) > c$ and $\widetilde{f}(\eta') > c$. But then $\widetilde{f}(\eta+1) > c$ and $\widetilde{f}(\eta'+1) > c$. Consequently, if the agents of degree η and η' are not connected, they would like to form a link and g would not be stable. Let $S = \{i \in N(g) | 1 < \eta_i < n-1\}$. All agents in S are linked by the above. The agents in $N \setminus S$ have degree $n-1$, they are linked to all agents. An agent in S is linked to all other agents in S and to agents in $N \setminus S$. But then an agent in S has degree $n-1$ contrary to hypothesis. Thus S is empty. Consequently, the stable networks g such that $l_{ij}(g-kl) \leq 2$ for all i,j and all kl have only nodes with degree $n-1$ and are thus complete graphs. ■

Corollary 1 *If $\widetilde{f}$ is increasing, then:*

- *no core periphery structure is stable*

– *no homogeneous island structure where each agent has at least two links to the other islands is stable.*

Having established these results, we will now return to the stability of the new structures with short diameters in the case where $\widetilde{f}$ is decreasing. Recall that by Proposition 2, the core-periphery structure in Proposition 2 with periphery degree η is PS if

– $\widetilde{f}(\eta+1) < c \leq \widetilde{f}(\eta)$
– $\widetilde{f}(n-1) + (n-1-\eta-(\eta-1)(\eta-2))[b(2,\eta)-b(3,\eta)] > c$

These conditions can only be satisfied if $\widetilde{f}$ is decreasing. When this is the case, we note that it is fairly easy to find a cost range such that the conditions in Proposition 2 are satisfied when the periphery degree η is small compared to n. To satisfy the first condition we must take a cost $c > \widetilde{f}(\eta+1) \geq \widetilde{f}(n-1)$. Therefore the link to the center is only maintained if the indirect benefit term, bounded below by $(n-1-\eta-(\eta-1)(\eta-2))[b(2,\eta)-b(3,\eta)]$ is sufficiently large. When the network size n is large and η is small compared to n ($\eta << \sqrt{n}$), this term is large unless $[b(2,\eta)-b(3,\eta)]$ is very small, that is unless there is hardly any decay. In other words, when $\widetilde{f}$ is decreasing, it is easy to find cost ranges with PS core-periphery structures in which each peripheral agent has a "local" neighborhood that is relatively small compared to the whole network.

It is more difficult to obtain a stable core-periphery structure when the periphery degree η is large. Indeed, suppose for example that $n/2+1 < \eta < n-1$. In this case the only benefit that is lost when a peripheral agent breaks with the center is the direct utility of the link to the center. When $n/2+1 < \eta < n-1$, the peripheral agent can reach the whole network at distance 2 without going through the center. Peripheral agent i has at least $n/2+1$ contacts other than the center. Suppose that there were some agent k that he could not reach at distance 2 through these contacts. But agent k has at least $n/2+1$ links that are not to the center. At least one of these must lead to a neighbor of i. The benefit of conserving the link to the center is then reduced to $b(1,n-1)-b(2,n-2) = \widetilde{f}(n-1) < \widetilde{f}(\eta) < c$ and so the core-periphery structure is not stable. It is quite natural that a core periphery structure where agents have very large peripheral neighborhoods cannot be stable. Indeed, the peripheral agents can then reach a large fraction of the network without going through the center which then becomes superfluous and cannot be maintained.

While there may not be any stable core periphery structure with a large periphery degree η, if η is sufficiently large, then we can always find a cost range in which there is a stable homogeneous island structure (Proposition 4) with degree η when $\widetilde{f}$ is decreasing.

Proposition 7 *For degrees η such that the structure in Proposition 4 exists, this structure is PS if $\widetilde{f}(\eta+1) < c \leq \widetilde{f}(\eta)$. Such a c always exists when $\widetilde{f}$ is decreasing.*

Proof: We revisit the conditions of Proposition 4. When $\widetilde{f}$ is decreasing, we can always find a cost satisfying $\widetilde{f}(\eta+1) < c \leq \widetilde{f}(\eta)$. When $\widetilde{f}$ is decreasing, the latter also implies that $f(d) > c$ for any $d < \eta$. ■

The propositions in this section highlight the importance of whether the function $\widetilde{f}$ is decreasing or increasing. This can in turn be related to the concavity/convexity of the benefit function with respect to degree. However, even without concavity/convexity of the latter, it is easy to compute directly the function $\widetilde{f}$ whose behavior allows us to pin down the stable structures with short diameters. Structures with short diameters other than the star and the complete network, such as the core-periphery network and the homogeneous degree network, cannot be stable when $\widetilde{f}$ is increasing. On the other hand, when $\widetilde{f}$ is decreasing, we always have a cost range where the homogeneous island structure is stable. A cost range where the core-periphery structure is stable should also exist for many reasonable benefit functions at least when the number of peripheral links is small compared to the whole network. This shows that our model exhibits new pairwise stable structures in some cost ranges under reasonable and not too demanding assumptions. Indeed, the function $\widetilde{f}$ will be decreasing if the benefit function is concave with respect to the degree, or even when it is convex in this variable, provided that decay is high enough.

We also observe that the comparison of link costs and the marginal benefit of linking to an agent of degree η, i.e. $\widetilde{f}(\eta)$ is important for determining whether a core-periphery or a homogeneous structure emerges. If $\widetilde{f}(\eta)$ exceeds the cost even for large degrees η, then it is difficult to maintain a core-periphery structure. The peripheral agents want to form many links in the periphery, but by doing so, they are able to circumvent the center which becomes superfluous and cannot be maintained. On the other hand, the homogeneous organization where all agents have an equal and fairly high but not maximal degree will be stable. Such a structure is not possible if the cost exceeds $\widetilde{f}(\eta)$ for a small degree. In this case, however, we can have a stable core-periphery structure with only a small number of peripheral links.

4.4 Pairwise stability for extreme values of the decay parameter

We complete our analysis of pairwise stability by considering the two extreme cases where decay is very large (i.e., δ is very small for the model given by (7)), or where decay is very small (i.e., δ is close to one for the model defined by (7)). We will show that in both of these cases, the degree-distance-based model can exhibit a very large number of PS architectures, which are only restricted by the fact that (most) nodes must have the same degree. When decay is large, the pairwise stable architectures include a number of disconnected structures with constraints on the degrees. These structures can be seen to coincide with those that are shown to be stable in Morrill (2011) (Proposition 2, p. 372). This is a natural since the benefits of a direct link in our model coincide with the benefits of a link in Morrill (2011) and that the indirect benefits can be neglected when decay is large.

The case of the benefit function where decay is very large can be expressed by the condition $b(1,k) >> b(2,k)$ for every k.

Proposition 8 *Let n be a fixed network size. Let $\epsilon > 0$. Then there exists $\underline{b} > 0$ such that for any function b with $b(2,1) < \underline{b}$ and any cost such that $b(1,r+1) + \epsilon < c < b(1,r) - \epsilon$, a network g satisfying the following properties is pairwise stable: in g, $n-k$ nodes, where $k \leq r$, have an identical degree r. The remaining k nodes are all linked to each other (such a network is what Morrill (2011) calls a maximal nearly k-regular network).*

Proof: Fix $\epsilon > 0$. Consider the maximal indirect benefit an agent can gain from a link. This benefit is bounded by $(n-2)b(2,1)$. Let $\underline{b} = \frac{\epsilon}{n-2}$. For any $0 < b(2,1) < \underline{b}$, the benefit of indirect links is inferior to ϵ. Basically, we can now neglect utility from indirect contacts. Let us establish that no pair of agents i, j in g can establish a mutually beneficial link. Let i be a node of degree $\eta_i = r$. Then $\widetilde{u}_j(g+ij) - \widetilde{u}_j(g) < b(1, r+1) + \epsilon - c < 0$. Therefore no agent wishes to form a link to an agent who already has degree r. Let $ij \in g$ with $\eta_i \leq r$ and $\eta_j \leq r$. Neither i nor j wishes to break this link: $\widetilde{u}_j(g-ij) - \widetilde{u}_j(g) < c + \epsilon - b(1, \eta_i) \leq c + \epsilon - b(1, r) < 0$. ∎

The family of PS networks described in Proposition 8 is very large and includes in particular all structures where agents have identical degrees, such as the circle or a generalized circle with agents linked to their m nearest neighbors. There is also an abundance of disconnected structures that satisfy the condition stated in Proposition 8.

One example is that of a number of disconnected "islands" of identical size.

Corollary 2 *Consider a network of size n. Let m be a divisor of n. The network consisting of n/m completely connected components of size m is PS under the conditions stated in Proposition 8.*

The class of PS networks described in Proposition 8 exists for some decay values and for some cost range for every network size. However, in order for these structures to appear, decay must be large, and all the more so when n is large, as we have $(n-2)b(2,1) < \epsilon < b(1,r) \leq b(1,1)$. The possible values of the cost for which these structures exist shrink the larger n and r are. In practice, the most likely decay is very large so that indirect benefits are almost negligible. The larger the size of the components, the smaller is the possible cost range. Among the possible structures, the one that we are most likely to see emerge in practice is thus that in which all agents have degree one or two, and the network size is not too large.

Let us now consider the other extreme case where decay is very small so that δ is close to one in the model defined by (7). In other words, we consider the benefit functions such that $b(l,k) \approx b(l',k)$ for all $k, l, l' \neq \infty$.

Proposition 9 *Let n be a fixed network size. Let $c > 0$. Then there exist $\epsilon > 0$ and b such that $|b(l,k) - b(l',k)| \leq \epsilon$ for all $k, l, l' \neq \infty$, and any network g satisfying the following properties is pairwise stable: g is minimally connected and satisfies $\min_{j \in N} b(1, \eta_j(g)) \geq c > (n-1)\epsilon$.*

Proof: Fix $c > 0$. What is the maximal benefit an agent can derive from an additional link in an already connected structure? The indirect benefit of link formation is bounded above by $(n-2)\max_{1 \leq x \leq n-1}(b(1,x) - b(n-1,x)) \leq (n-2)\epsilon$. Since g is connected, any additional link provides a utility $\widetilde{u}_i(g+ij) - \widetilde{u}_i(g) \leq b(1, \eta_j+1) - b(n-1, \eta_j) + (n-2)\epsilon - c \leq (n-2)\epsilon + b(1, \eta_j+1) - b(n-1, \eta_j+1) + b(n-1, \eta_j+1) - b(n-1, \eta_j) - c \leq (n-1)\epsilon - c < 0$. Let us establish that no agent wants to remove a link. If $ij \in g$, then neither i nor j wishes to break this link. Indeed, because g is minimally connected, i and j are not in the same connected component in $g - ij$. Therefore $\widetilde{u}_i(g-ij) - \widetilde{u}_i(g) < c - b(1, \eta_j) < 0$. ∎

De Jaegher and Kamphorst (2013) also study a model where agents' payoffs are based on their access to information in a setting with small decay. Information is defined as the sum of all decayed paths to indirect contacts, i.e. exactly the benefit term in the connections model. Payoffs differ from those of Jackson and Wolinsky (1996) because agents apply a possibly non-linear function to evaluate this aggregate quantity. Our model differs from Jackson and Wolinsky (1996) because we impose a "penalty" on the value received from each indirect contact *before* aggregating the value of all indirect contacts by addition. Because of these differences, our setting and that of De Jaegher and Kamphorst (2013) are not directly comparable. However, in both cases, the small decay assumption leads to stable structures that are minimally connected, reflecting the low benefit of forming a link to someone who is already in the same connected component. We obtain this result when the decay parameter approaches 1, whereas De Jaegher and Kamphorst (2013) assume small but not vanishing decay.

5 Efficiency in the model

Next, we analyze strong efficiency (SE) in the model, also sometimes referred to as efficiency in the paper. In the degree-distance-based model, there is a much wider range of possible SE architectures than in the Jackson and Wolinsky model. While in the latter, only $g^{\emptyset}$, g^{*} and g^{N} can be SE, in our model additional structures can be SE for some parameters. To see this immediately, consider the 3-player example presented in Section 4.2. Figure 7 shows all strongly efficient networks in this model given by (7).

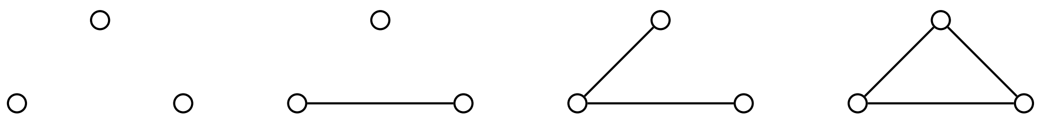

Fig. 7. Unique SE networks in the degree-distance-based connections model given by (7) (from left to right): (i) ($0 < \delta < \frac{1}{3}$ and $2c > \delta$) or ($\frac{1}{3} \leq \delta < 1$ and $2c > \frac{1}{2}\delta^2 + \frac{5}{6}\delta$), (ii) $0 < \delta < \frac{1}{3}$ and $\delta^2 + \frac{2}{3}\delta < 2c < \delta$, (iii) ($0 < \delta < \frac{1}{3}$, $\frac{1}{3}\delta - \delta^2 < 2c < \delta^2 + \frac{2}{3}\delta$) or ($\frac{1}{3} \leq \delta < 1$, $2c < \frac{1}{2}\delta^2 + \frac{5}{6}\delta$), (iv) $2c < \frac{1}{3}\delta - \delta^2$

In this simple 3-player example we can already see an additional difference between our model and the original connections model. In the latter, the cost ranges of PS and SE for the complete network coincide. In the degree-distance-based model defined by (7), g^N can be PS but not efficient: for $n = 3$ it is the case for $\frac{1}{3}\delta - \delta^2 < 2c < \frac{2}{3}\delta - \delta^2$. Similarly to Jackson and Wolinsky, we exhibit the contradiction between stability and efficiency in the higher cost ranges: the empty network is PS but not SE for $\delta < 2c < \frac{1}{2}\delta^2 + \frac{5}{6}\delta$ and $\frac{1}{3} < \delta < 1$. The structure which could neither be PS nor SE in the original connections model – the network containing one link and one isolated player – is PS and SE in our model in the same cost range.

Another interesting observation in the model given by (7) and $n = 4$ is for instance that the line is the unique SE network if $0 < \delta < \frac{4-\sqrt{13}}{6}$ and $\frac{1}{3}\delta - \frac{1}{3}\delta^2 - \delta^3 < 2c < \frac{1}{3}\delta + \frac{5}{3}\delta^2 + \delta^3$; for the calculations, see the Appendix.

After presenting these examples, we turn to the theoretical analysis. Contrary to the original connections model, disconnected networks may be pairwise stable. Let us show that they can also be efficient.

Proposition 10 *Let n be even and fixed, and $\epsilon > 0$. There exists $\underline{b} > 0$ such that for any function b with $b(2,1) < \underline{b}$, the network described in Proposition 8, consisting of $n/2$ disjoint completely connected components with $m = 2$ is uniquely efficient in the cost range $\frac{b(1,1)+b(1,2)}{2} + \epsilon < c < b(1,1) - \epsilon$.*

Proof: Fix $\epsilon > 0$. Consider the maximal indirect benefit an agent can gain from a link. This benefit is bounded by $(n-2)b(2,1)$. Thus the total social utility from indirect links is bounded by $n(n-2)b(2,1)$. Let $\underline{b} = \frac{\epsilon}{n(n-2)}$. For any $0 < b(2,1) < \underline{b}$, the total social utility of indirect links is inferior to ϵ. Basically, we can now neglect utility from indirect contacts. Note that if there are at least two nodes with no link, forming a link between them increases total utility since $c < b(1,1)$. Assume now that a network contains some node i such that $\eta_i = k \geq 2$. Let ij belong to the network. Let us show that removing ij is efficiency improving. Let $\eta_j \geq 1$. Since indirect benefits are negligible, only i and j lose by removing the link ij. This loss is $b(1,\eta_i) + b(1,\eta_j) \leq b(1,2) + b(1,1) < 2c$. Therefore removing the link of an agent with degree greater than one is efficiency improving. In this parameter range, an efficient network is such that each agent has exactly degree 1. This is achieved uniquely by a network consisting of disjoint connected components of size two. ∎

When decay is very large, benefits from indirect contacts are negligible and the negative impact of an increased degree dominates. Thus it is not socially desirable to connect two components.

Proposition 11 shows that when network size is large, the complete network is not strongly efficient when it is stable.

Proposition 11 *Whenever $(n-1)\left(b(1,n-2) - b(1,n-1)\right) > b(1,n-2) - b(2,n-2)$, the complete network is not strongly efficient for any cost $c > 0$. For the model defined by (7), g^N is not SE whenever $n > \frac{1}{\delta}$. In particular, the complete network is not strongly efficient when it is uniquely PS.*

Proof: We note that the total link cost is always greater in g^N than in $g^N - ij$ when $c > 0$. Assume $c = 0$. Let us consider the difference in total utility between g^N and $g^N - ij$, that is $\sum_{i=1}^{n}\left(\widetilde{u}_i(g^N) - \widetilde{u}_i(g^N - ij)\right)$. For agents i, j the utility loss is $b(1,n-1) - b(2,n-2)$. The remaining $n-2$ agents gain from the connectivity decrease of their direct neighbors i and j, $2(n-2)\left(b(1,n-2) - b(1,n-1)\right)$. In total, the change in social utility is $\sum_{i=1}^{n}\left(\widetilde{u}_i(g^N - ij) - \widetilde{u}_i(g^N)\right) = 2(n-2)\left(b(1,n-2) - b(1,n-1)\right) - 2b(1,n-1) + 2b(2,n-2) = 2(n-1)\left(b(1,n-2) - b(1,n-1)\right) + 2\left(b(2,n-2) - b(1,n-2)\right)$. This quantity is positive whenever $(n-1)\left(b(1,n-2) - b(1,n-1)\right) > b(1,n-2) - b(2,n-2)$ meaning that the complete network g^N is not strongly efficient. Except for very large decay (when $b(1,n-2) >> b(2,n-2)$), the complete network is inefficient even when n is rather small. Solving $(n-1)\left(b(1,n-2) - b(1,n-1)\right) > b(1,n-2) - b(2,n-2)$ for the model defined by (7) leads equivalently to the condition $n > \frac{1}{\delta}$. ∎

Proposition 12 *Let g be a network in which $l_{ij}(g) \leq 2$ for all i, j and let kl be a link such that $l_{ij}(g - kl) \leq 2$ for all i, j. Then:*

(i) $g-kl$ has higher overall utility than g for any $c>0$ when n is such that $2(n-2) > \frac{K}{min(\alpha,\beta)}$ where $K = b(1,\eta_k) - b(2,\eta_k-1) + b(1,\eta_l) - b(2,\eta_l-1)$, $\alpha =: min_{1<l<n}[b(1,l-1) - b(1,l)] > 0$ and $\beta =: min_{1<l<n}[b(2,l-1) - b(2,l)] > 0$

(ii) In particular, for the model defined by (7), $g-kl$ is strictly more efficient than g for any $c>0$ if $\delta > 1/n$.

Corollary 3 *The windmill (presented in Proposition 2), the complete graph or the multiple island model (presented in Proposition 4 with $l>1$) are never efficient even when they are PS, if n satisfies the condition in Proposition 12.*

Indeed, it is readily verified that these structures contain a link whose removal conserves a maximal network diameter of 2.

Proof: (of Proposition 12) Consider two nodes k,l, such that the maximal distance in $g-kl$ is still 2. We must have $1<\eta_k \leq n-1$ and $1<\eta_l \leq n-1$. Let us show that overall utility increases when this link is removed.

First consider the change in utility for k and l: $\widetilde{u}_k(g) - \widetilde{u}_k(g-kl) + \widetilde{u}_l(g) - \widetilde{u}_l(g-kl) = b(1,\eta_k) - b(2,\eta_k-1) - c + b(1,\eta_l) - b(2,\eta_l-1) - c$. This quantity is positive when g is pairwise stable, since k and l have an incentive to maintain the link. It is bounded by K.

The presence of the link kl has a negative impact on all other agents. Agent k has degree η_k. Thus he has η_k-1 direct neighbors, excluding l already accounted for above. Besides k himself and his η_k direct neighbors, the remaining $n-1-\eta_k$ agents are at distance 2. The utility loss for these agents is $(\eta_k-1)(b(1,\eta_k-1) - b(1,\eta_k)) + (n-1-\eta_k)(b(2,\eta_k-1) - b(2,\eta_k))$ and similarly for agent l, replacing η_k by η_l.

We can now compare the overall utility of g and $g-kl$. It is $\sum_{i=1}^n (\widetilde{u}_i(g) - \widetilde{u}_i(g-kl)) = b(1,\eta_k) - b(2,\eta_k-1) + b(1,\eta_l) - b(2,\eta_l-1) - 2c - (\eta_k-1)(b(1,\eta_k-1) - b(1,\eta_k)) - (n-1-\eta_k)(b(2,\eta_k-1) - b(2,\eta_k)) - (\eta_l-1)(b(1,\eta_l-1) - b(1,\eta_l)) - (n-1-\eta_l)(b(2,\eta_l-1) - b(2,\eta_l))$

Define $\alpha =: min_{1<l<n}[b(1,l-1) - b(1,l)] > 0$ and $\beta =: min_{1<l<n}[b(2,l-1) - b(2,l)] > 0$. $\sum_{i=1}^n (\widetilde{u}_i(g) - \widetilde{u}_i(g-kl)) \leq K - 2(n-2)min(\alpha,\beta) - 2c < K - 2(n-2)min(\alpha,\beta)$, which is negative provided $2(n-2) > \frac{K}{min(\alpha,\beta)}$.

In particular, for the model defined by (7), $\sum_{i=1}^n (u_i(g) - u_i(g-kl))$ is negative if $F(\eta_k) + F(\eta_l) < 0$ with $F(\eta_k) = \frac{\delta}{\eta_k+1} - \frac{\delta^2}{\eta_k} - (\eta_k-1)\left[\frac{\delta}{\eta_k} - \frac{\delta}{\eta_k+1}\right] - (n-1-\eta_k)\left[\frac{\delta^2}{\eta_k} - \frac{\delta^2}{\eta_k+1}\right]$ $F(\eta_k) < 0 \iff \eta_k - \delta(\eta_k+1) - (\eta_k-1) - (n-1-\eta_k)\delta < 0 \iff 1 < n\delta$. Similarly, $F(\eta_l) < 0$ if and only if $n > \frac{1}{\delta}$. Thus, whenever $n > \frac{1}{\delta}$ the network g is not efficient. ■

This result is easy to understand. In a network in which all agents are at distance at most 2 when kl is removed, the link kl benefits only agents k and l themselves and exerts a negative externality on a large number of agents at distance 2. Provided n is large, this outweighs the positive effects of the link.

Next, we focus on the conditions for efficiency of the star. Let us define the function

$$h(m) = mb(1,m) + (n-1-m)b(2,m) \tag{13}$$

This function represents an upper bound of the social utility that an agent with degree m provides to others. This upper bound is attained for example for a star, but generally it is not attained. For the model defined by (7), we have $h(m) = m\frac{\delta}{1+m} + (n-1-m)\frac{\delta^2}{1+m}$.

We assume the following conditions:

Condition 1 *Function h defined by (13) is decreasing.*

Condition 2 *For any $k > 2$, let $a + d = k$ and $a^{'} + d^{'} = k$. Then if $|a - d| \geq |a^{'} - d^{'}|$ then $h(a) + h(d) \geq h(a^{'}) + h(d^{'})$.*

Condition 2 means that when keeping fixed the sum of degrees of two agents, the total maximal social utility provided by the two agents, as measured by the function h is greater if the agents have dissimilar degrees. Note that Conditions 1 and 2 are satisfied for the model defined by (7), when $n > \frac{1}{\delta}$; for the proof, see the Appendix.

Condition 1 is not demanding and holds easily if the network size n is large. Condition 2 seems to have a convexity flavor. One agent with high degree and one agent with low degree provide more social utility (or at least the upper bound given by h is greater) than two agents with intermediate degrees. In fact the proposition below shows under some conditions that Condition 2 cannot hold if b is concave with respect to degree. However, convexity is not sufficient to guarantee Condition 2 because a convex function whose behavior is very close to that of a linear function would not satisfy it.

Proposition 13 *Suppose that for any degree η, $b(2, \eta) = \delta b(1, \eta)$ for some $\delta \leq 1$ (decay independent of degree) and that $b(1, \eta)$ is differentiable. Then Condition 2 never holds if $b(1, \eta)$ is concave. Moreover Condition 2 does not hold for all convex $b(1, \eta)$.*

This proposition is proved in the appendix.

Proposition 14 *Let g be a connected structure. Whenever h given by (13) satisfies Conditions 1 and 2, we have $v(g) \leq v(g^*)$.*

Proof: We will prove this proposition in two steps.

Step 1: *First we show that $v(g) \leq v(g^*)$ for any minimally connected structure g.*

Let g be a minimally connected structure. It is thus characterized by m, J_1 and S in Lemma 1 (presented in the Appendix), the degrees of the nodes in S and the distances between all pairs of nodes. By Lemma 1, all nodes in J_1 have degree one. There are m nodes in S whose degrees are $(2 + \alpha_i)_{i=1}^{i=m}$ (without loss of generality we let the m nodes in S be $1, 2, ..., m$). The remaining $n - j_1 - m$ nodes have degree 2. Therefore by Lemma 2 (presented in the Appendix), we have
$v(g) \leq \sum_{i=1}^{i=n} h(\eta_i(g)) = j_1 h(1) + \sum_{i=1}^{i=m} h(2 + \alpha_i) + (n - j_1 - m)h(2)$.
We apply Condition 2 to obtain $h(\alpha_i + \alpha_j + 2) + h(2) \geq h(\alpha_i + 2) + h(\alpha_j + 2)$. Thus we have $\sum_{i=1}^{i=m} h(\alpha_i + 2) + (n - m - j_1)h(2) = \sum_{i=1}^{i=m-2} h(\alpha_i + 2) + (n - m - j_1)h(2) + h(\alpha_{m-1} + 2) + h(\alpha_m + 2) \leq \sum_{i=1}^{i=m-2} h(\alpha_i + 2) + (n - m - j_1)h(2) + h(\alpha_{m-1} + \alpha_m + 2) + h(2) = \sum_{i=1}^{i=m-2} h(\alpha_i + 2) + h(\alpha_{m-1} + \alpha_m + 2) + (n - m - j_1 + 1)h(2) \leq .... \leq h(\sum_{i=1}^{i=m} \alpha_i + 2) + (n - j_1 - 1)h(2)$.
Then, applying again repeatedly Condition 2, $v(g) \leq j_1 h(1) + h(\sum_{i=1}^{i=m} \alpha_i + 2) + (n - j_1 - 1)h(2) = j_1 h(1) + h(j_1) + (n - j_1 - 1)h(2) = j_1 h(1) + (n - j_1 - 2)h(2) + h(2) + h(j_1) \leq j_1 h(1) + (n - j_1 - 2)h(2) + h(1) + h(j_1 + 1) = (j_1 + 1)h(1) + h(j_1 + 1) + (n - j_1 - 2)h(2) \leq ... \leq (n - 1)h(1) + h(n - 1) = v(g^*)$.

Step 2: *Next, we prove that $v(g) \leq v(g^*)$ for any connected structure g.*

As we have shown, any minimally connected structure g has a degree sequence $(\eta_i)_{i=1}^{i=n}$ such that $v(g) \leq \sum_{i=1}^{i=n} h(\eta_i) \leq v(g^*)$. Any connected structure g_K is a superset of a minimally connected network g. Let $g \subset g_K$ and let $\mu_i = \eta_i(g_K) - \eta_i(g) \geq 0$. By Lemma 2 we have $v(g_K) \leq \sum_{i=1}^{i=n} h(\eta_i(g_K))$. We will show that $\sum_{i=1}^{i=n} h(\eta_i(g_K)) \leq \sum_{i=1}^{i=n} h(\eta_i(g))$. Let us consider h being decreasing. For the model given by (7), we verify that for all $m > 1$, $h(m+1) - h(m) \leq 0 \iff \delta \geq \frac{1}{n}$. We use this now to show successively that: $\sum_{i=1}^{i=n} h(\eta_i(g_K)) = \sum_{i=1}^{i=n} h(\eta_i(g) + \mu_i) \leq \sum_{i=1}^{i=n} h(\eta_i(g) + (\mu_i - 1)) \leq \sum_{i=1}^{i=n} h(\eta_i(g) + (\mu_i - 2)) \leq .... \leq \sum_{i=1}^{i=n} h(\eta_i(g))$. Since $\sum_{i=1}^{i=n} h(\eta_i) \leq v(g^*)$, we conclude that $v(g_K) \leq \sum_{i=1}^{i=n} h(\eta_i) \leq v(g^*)$. ■

Consequently, for the model defined by (7), $v(g) \leq v(g^*)$ whenever $n > \frac{1}{\delta}$. We now show that under some fairly weak assumptions on the payoffs, we also have $v(g^*) \geq v(g)$ for any disconnected network g. Under these conditions, the star will then be efficient.

Proposition 15 *Let g_1^* and g_2^* be two disjoint stars with centers i and j. Whenever $(n_k - 1)[b(1, n_k) + b(2, 1) - b(1, n_k - 1)] \geq c$ for $k = 1, 2$ (sufficient but not necessary condition), where n_k is the cardinality of g_k^*, $v(g_1^* \cup g_2^* + ij) \geq v(g_1^* \cup g_2^*)$. In particular, this cost range exists when $b(1, n_k) + b(2, 1) > b(1, n_k - 1)$.*

Proof: Under the assumptions in Proposition 14, the value of a star is not smaller than the value of any connected structure. Let g be a disconnected network. Then $v(g)$ is maximized when g is the union of star components. Let us show that under some weak conditions, social utility is increased by connecting two stars.

Indeed, let i and j be the centers of two stars of size n_1 and n_2, with $2 \leq n_1 \leq n$ and $2 \leq n_2 \leq n$. If we add a link between the centers, then the change in utility is $(n_1 - 1)[b(2, n_2) + (n_2 - 1)b(3, 1) + b(1, n_1) - b(1, n_1 - 1)] + b(1, n_2) + (n_2 - 1)b(2, 1) + (n_2 - 1)[b(2, n_1) + (n_1 - 1)b(3, 1) + b(1, n_2) - b(1, n_2 - 1)] + b(1, n_1) + (n_1 - 1)b(2, 1) - 2c \geq (n_1 - 1)[b(2, n_2) + b(1, n_1) + b(2, 1) - b(1, n_1 - 1)] + (n_2 - 1)[b(2, n_1) + b(1, n_2) + b(2, 1) - b(1, n_2 - 1)] - 2c$

This quantity is positive under a fairly weak condition: it is sufficient that decay with distance is not too great and utility decrease with respect to the neighbor's degree is not too great: $(n_k - 1)[b(1, n_k) + b(2, 1) - b(1, n_k - 1)] \geq c$ for $k = 1, 2$.

If connecting two disjoint stars is efficiency improving, the efficient network cannot be disconnected and so under Conditions 1 and 2, the star is efficient. ■

The condition in the above proposition is sufficient but not necessary and can be improved. However, the star is not uniquely efficient. The complete graph, the empty one, and also the line or a disconnected graph with connected components of size two can all be efficient for some choices of b. The results we have are sufficient to show that the star is uniquely stable under conditions which are compatible with the (in some cases unique) pairwise stability of other structures than the star.

From Propositions 14 and 15, and adding condition that $b(1, 1) > c$ (which ensures that the empty network is not efficient), we obtain the following.

Proposition 16 *Let the benefit function b satisfy Conditions 1 and 2, $(\widetilde{n} - 1)[b(1, \widetilde{n}) + b(2, 1) - b(1, \widetilde{n} - 1)] \geq c$ for all $2 \leq \widetilde{n} \leq n$ and let $b(1, 1) > c$. Then the star is efficient, and is uniquely efficient whenever a strict inequality holds in Condition 2.*

Having established the efficiency of the star under some conditions, we can now compare with the pairwise stable structures characterized in the previous section. The conditions for efficiency of the star can be compatible with the stability conditions of the complete network or with that of the windmill network (Propositions 1(iv) and 2). This is most readily verified by checking the respective conditions for the function (7). Indeed, the assumptions under which we proved the star to be efficient exclude concavity of the benefit function with respect to degree, but the new structures with short diameters could be stable both when the benefit function is concave and when it is convex in degree. This implies the existence of benefit functions and cost ranges verifying our general assumptions for which the star is efficient but not pairwise stable, and the pairwise stable or even uniquely pairwise stable network is not efficient. We have already seen that for large n the windmill or complete network are not efficient in their stability region. The result about the efficiency of the star also shows that the efficient network can be strictly contained in a (or the, in the case of uniqueness) pairwise stable network. This implies that our model can give rise to overconnectivity in the strong sense defined by Buechel and Hellmann (2012), which could never occur in the original connections model.

6 Conclusion

In this paper, we analyzed network formation in the presence of negative externalities in a model that combined the presence of indirect benefits and a penalty resulting from the connectivity of direct and indirect neighbors. Our analysis focused mainly on the case of structures with short diameters but also considered cases with extreme levels of decay. It would be interesting but more challenging to extend it beyond these cases. While remaining in the framework with global positive spillovers, we could also consider somewhat different models that capture other types of negative externalities. As we discussed in the introduction, the model we proposed here is a good fit for a situation that we could see as a "generalized" co-author model where knowledge spills over from more distant parts of the network. We can also think of situations where benefits could spill over from distant neighbors but be reduced by overall connectivity. One might see a link as a consumer good whose value is based on its rarity and which thus decreases the more common or widespread it is. We could also gear the model more specifically towards the competition for information, by letting payoffs depend on the number of informed agents in a communication chain. Finally, we note that in the model we considered, as well as in the potential extensions, we are likely to have a high multiplicity of equilibria, suggesting that the use of stronger stability concepts, e.g., strongly stable networks (those which are stable against changes in links by any coalition of individuals; see Jackson and van den Nouweland (2005)) could be helpful for equilibrium selection.

Appendix

Line is the unique SE network for $n = 4$, model (7) and some δ, c

Proof: Let $n = 4$. Let $U^{(k)} = \sum_{i=1}^{4} u_i(g^{(k)})$, where $k \in \mathbb{N}$ and $1 \leq k \leq 8$. Similarly, the sum of the players' utilities for the line, the empty graph and the complete graph

is denoted by U^L, $U^\emptyset$ and U^N, respectively. We will determine the parameters δ and c under which the line g^L is the unique SE network. We have: $U^L = \delta^3 + \frac{5}{3}\delta^2 + \frac{7}{3}\delta - 6c$ $U^\emptyset = 0$, $U^N = 3\delta - 12c$, for the graph with one link: $U^{(1)} = \delta - 2c$
for the graphs with two links: $U^{(2)} = 2\delta - 4c$, $U^{(3)} = \delta^2 + \frac{5}{3}\delta - 4c$
for the graphs with 3 links (different from g^L): $U^{(4)} = 3\delta^2 + \frac{9}{4}\delta - 6c$ (star), $U^{(5)} = 2\delta - 6c$
for the graphs with 4 links: $U^{(6)} = \frac{4}{3}\delta^2 + \frac{8}{3}\delta - 8c$, $U^{(7)} = \frac{5}{3}\delta^2 + \frac{31}{12}\delta - 8c$
for the graph with 5 links: $U^{(8)} = \frac{2}{3}\delta^2 + \frac{17}{6}\delta - 10c$.
$U^L > U^{(4)}$ iff $\delta^2 - \frac{4}{3}\delta + \frac{1}{12} > 0$ iff $0 < \delta < \frac{4-\sqrt{13}}{6}$. We solve: $U^L > U^{(2)}$ and $U^L > U^{(6)}$ and $0 < \delta < \frac{4-\sqrt{13}}{6}$, which gives $0 < \delta < \frac{4-\sqrt{13}}{6}$ and $\frac{1}{3}\delta - \frac{1}{3}\delta^2 - \delta^3 < 2c < \frac{1}{3}\delta + \frac{5}{3}\delta^2 + \delta^3$. For such δ and c, we have, in particular, $\delta > 2c$. Hence, $U^L > U^{(1)}$, $U^L > U^\emptyset$, and also $U^L > U^{(3)}$, $U^L > U^{(5)}$, $U^L > U^{(8)}$, $U^L > U^{(7)}$, $U^L > U^N$. ■

Conditions 1 and 2 are satisfied for the model (7), when $n > \frac{1}{\delta}$

Proof: Consider $h(m) = m\frac{\delta}{1+m} + (n-1-m)\frac{\delta^2}{1+m}$. We have $h'(m) = \frac{\delta(1-\delta n)}{(1+m)^2} < 0$ for $n > \frac{1}{\delta}$. Consider a, d such that $a + d = k$, and therefore $d = k - a$. For the model defined by (7), we have $h(a) + h(d) = h(a) + h(k-a) = \frac{\delta a}{1+a} + (n-1-a)\frac{\delta^2}{1+a} + (k-a)\frac{\delta}{1+k-a} + (n-1-(k-a))\frac{\delta^2}{1+k-a} = \frac{(1+a)(\delta-\delta^2)}{1+a} + \frac{(1+k-a)(\delta-\delta^2)}{1+k-a} + \delta\left[\frac{n\delta-1}{1+a} + \frac{n\delta-1}{1+k-a}\right]$. Thus we can write $h(a) + h(d) = c(\delta) + \delta G(a)$, where $G(a) =: \frac{n\delta-1}{1+a} + \frac{n\delta-1}{1+k-a}$. The derivative of this function is $G'(a) = \frac{1-n\delta}{(1+a)^2} + \frac{n\delta-1}{(1+k-a)^2}$ which is zero at $a = k/2$. Moreover, we can show that this zero corresponds to a min provided that $n\delta > 1$. Indeed, when this is the case, $G(a) \geq 0$ and $G(0) = (n\delta-1)\left(1 + \frac{1}{1+k}\right) \geq G(k/2) = (n\delta-1)\left(\frac{2}{1+k/2}\right)$ for every $k \geq 2$. We also note that by symmetry, $G(a) = G(k-a)$. This implies that $G(a)$ is decreasing on $[0, k/2]$ and increasing on $[k/2, k]$. Since $|a-d| = |a-(k-a)| = |2a-k|$, if $|a-d| > |a'-d'|$, then $|a-k/2| > |a'-k/2|$, which implies $G(a) > G(a')$, which implies $h(a) + h(d) > h(a') + h(d')$. ■

Proof of Proposition 13

Proof: Condition 2 holds if and only if $h(a)+h(d) = h(a)+h(k-a) =: H(a)$ is decreasing on $[0, k/2]$. We have $b(2,\eta) = \delta b(1,\eta)$ for some $\delta \leq 1$. We compute $H'(a)$ to find $H'(a) = (1-\delta)[b(1,a)-b(1,k-a)]+\delta(n-1)[b'(1,a)-b'(1,k-a)]+(1-\delta)[ab'(1,a)-(k-a)b'(1,k-a)]$. Now $a \leq k-a$. Consequently if $b(1,\eta)$ is concave in degree, $b'(1,k-a) < b'(1,a) < 0$ and $H'(a) > 0$. Thus Condition 2 is not compatible with concavity of b with respect to degree. We can also see that in the limit case between concavity and convexity where $b(1,\eta)$ is linear in degree, we would have $H'(a) \geq (1-\delta)[b(1,a) - b(1,k-a)] > 0$. Thus convexity is not sufficient to guarantee Condition 2. ■

Lemma 1 *Let g be a minimally connected network of size $n > 3$. Let J_1 be the set of nodes of degree 1 and j_1 the number of elements in this set. Then, whenever $j_1 > 2$, there exists a set S containing $1 \leq m \leq j_1$ nodes such that for all $i \in S$, $\eta_i \geq 3$ and $\sum_{i\in S}\alpha_i = j_1 - 2$, with $\alpha_i = \eta_i - 2$.*

Proof: We prove this by induction on the network size. Any minimally connected structure of size $n+1$ can be obtained by adding one node $n+1$ and one link between $n+1$ and some $j < n+1$ in a minimally connected network of size n. Suppose that g_n $(n \geq 4)$ verifies the induction hypothesis. If $S(g_n) = \emptyset$, g_n is a line. If we link $n+1$ to a node of degree 1, $j_1(g_{n+1}) = j_1(g_n) = 2$ and $S(g_{n+1})$ is still empty. If we add a link between $n+1$ and a node of degree 2, there will be 3 nodes with degree 1 and 1 node of degree 3. Thus $j_1(g_{n+1}) = 3$, and $\sum_{i \in S(g_{n+1})} \alpha_i(g_{n+1}) = 1 = j_1(g_{n+1}) - 2$. If g_n verifies the induction hypothesis and $S(g_n) \neq \emptyset$, there are several possibilities. Either the link from $n+1$ goes to a node j such that $\eta_j(g_n) = 1$. Then the number of nodes with degree one does not change so $j_1(g_{n+1}) = j_1(g_n)$. The number of nodes in S does not change and $\eta_j(g_{n+1}) = 2$. Thus $\sum_{i \in S(g_{n+1})} \alpha_i(g_{n+1}) = \sum_{i \in S(g_n)} \alpha_i(g_n) = j_1(g_n) - 2 = j_1(g_{n+1}) - 2$. If $\eta_j(g_n) = 2$ then $\eta_j(g_{n+1}) = 3$ and $card(S(g_{n+1})) = card(S(g_n)) + 1$ and $j_1(g_{n+1}) = j_1(g_n) + 1$. Thus $\sum_{i \in S(g_{n+1})} \alpha_i(g_{n+1}) = \sum_{i \in S(g_n)} \alpha_i(g_n) + \alpha_j(g_{n+1}) = j_1(g_n) - 2 + 1 = j_1(g_{n+1}) - 2$. Finally, if $n+1$ links to j such that $\eta_j(g_n) > 2$, then $j_1(g_{n+1}) = j_1(g_n) + 1$, $card(S(g_n)) = card(S(g_{n+1}))$ and $\eta_j(g_{n+1}) = \eta_j(g_n) + 1$. Thus $\sum_{i \in S(g_{n+1})} \alpha_i(g_{n+1}) = \sum_{i \in S(g_n)} \alpha_i(g_{n+1}) = \sum_{i \in S(g_n)} \alpha_i(g_n) + 1 = j_1(g_n) - 2 + 1 = j_1(g_{n+1}) - 2$. This concludes the proof of the induction step. The induction hypothesis is verified when $n = 4$. There are two minimally connected structures: a line and a star. In a line, there are two elements of degree 1, thus $j_1 = 2$ and $S = \emptyset$. In a star with $n = 4$, $j_1 = 3$, S consists of the center with degree 3 and indeed $\sum_{i \in S(g_n)} \alpha_i(g_n) = 3 - 2 = j_1(g_n) - 2$. ■

Lemma 2 *Let g be a network with degree sequence $(\eta_i(g))_{i=1}^{i=n}$. Then the value of g is $v(g) \leq \sum_{i=1}^{i=n} h(\eta_i(g))$.*

Proof: The $\eta_i(g)$ immediate neighbors of i derive the utility $b(1, \eta_i(g))$ from the link to i. The remaining $n - 1 - \eta_i(g)$ nodes are at distance at least 2 from i and therefore the utility obtained from i is bounded by $b(2, \eta_i(g))$. ■

Bibliography

P. Billand, C. Bravard, and S. Sarangi. A note on local spillovers, convexity and the strategic substitutes property in networks. *Theory and Decision*, 2012. Forthcoming.

F. Bloch and M. O. Jackson. The formation of networks with transfers among players. *Journal of Economic Theory*, 133(1):83–110, 2007.

B. Buechel and T. Hellmann. Under-connected and over-connected networks: the role of externalities in strategic network formation. *Review of Economic Design*, 16:71–87, 2012.

K. De Jaegher and J. J. A. Kamphorst. Two-way flow networks with small decay. Working Paper, January, 2013.

S. Goyal and S. Joshi. Unequal connections. *International Journal of Game Theory*, 34: 319–349, 2006.

H. Haller. Network extension. *Mathematical Social Sciences*, 64:166–172, 2012.

M. O. Jackson. *Social and Economic Networks*. Princeton University Press, 2008.

M. O. Jackson and A. van den Nouweland. Strongly stable networks. *Games and Economic Behavior*, 51:420–444, 2005.

M. O. Jackson and A. Wolinsky. A strategic model of social and economic networks. *Journal of Economic Theory*, 71:44–74, 1996.

T. Morrill. Network formation under negative degree-based externalities. *International Journal of Game Theory*, 40:367–385, 2011.